On the minimum number of unitaries needed to describe a random-unitary channel

Francesco Buscemi[*]

ERATO-SORST Quantum Computation and Information Project,
Japan Science and Technology Agency,
Daini Hongo White Bldg. 201, 5-28-3 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

August 28th, 2006

Abstract

We provide, in an extremely simple way, an upper bound to the minimum number of unitary operators describing a general random-unitary channel.

1 Introduction

A channel $\mathcal{E}$—i. e. a completely positive trace-preserving map—acting on density matrices ρ defined on a finite dimensional input Hilbert space $\mathscr{H}$ (for sake of simplicity, we consider here channels with equal input and output Hilbert spaces; the generalization is straightforward) is called *random-unitary* if it admits a Kraus representation [1] as

$$\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger, \tag{1}$$

where p_i are probabilities and U_i are unitary operators. This definition involves an existential quantifier, and there is no known constructive algorithm

[*] email: `buscemi@qci.jst.go.jp`

to check whether a given channel is random-unitary or not. Only the necessary condition of being *unital*, that is, of preserving the identity matrix, $\mathcal{E}(\mathbf{1}) = \mathbf{1}$, holds[1].

Nonetheless, random-unitary channels play a very special *physical* role among all possible evolutions that an open quantum system can undergo [3]. In fact, Gregoratti and Werner [4] proved that they are the only irreversible channels that can be perfectly corrected using, as the only side-resource, classical information extracted from the environment. This property is actually sufficient and necessary for a channel to be random-unitary and one would prefer to adopt this one as the *physical and operational definition* of random-unitary channels. The idea of using the environment as a resource then initiated investigations about *environment-assisted capacities* for quantum channels [5, 6, 7]. In Ref. [8] the problem also of quantifying the amount of classical information needed to perfectly correct a random-unitary channel was raised for the first time in the case of decohering evolutions. In fact, for a given random-unitary channel, the form (1) is highly non-unique and the Shannon entropy $H(p_i)$ of the probability distribution weighing the unitaries U_i can be "artificially" made as large as desired. Consequently, in order to derive sensible information-theoretic relations regarding the information dynamics in a random-unitary evolution, one has to single out the random-unitary Kraus representation minimizing $H(p_i)$.

In the present paper we derive an upper bound for the minimal number of unitary operators needed in Eq. (1), thus providing also a bound to the amount $H(p_i)$ of classical information needed to be extracted from the environment in order to invert the random-unitary evolution. Our bound, proved for generic dimension, does not catch the peculiar geometry that bistochastic qubit channels enjoy: For $d = 2$, it is provably non tight. However, the qubit case is completely understood and all random-unitary qubit maps have already been explicitly characterized (see, e. g. Ref. [9]). In this sense, a bound for the qubit case is completely useless. On the contrary, as soon as one leaves the two-dimensional world, already for $d = 3$, the bound we provide is generally non trivial.

[1] For two-dimensional systems, a channel is random-unitary if and only if it is unital. For higher dimensional systems, if a channel is random-unitary it is also unital, but the converse does not hold [2].

2 Properties of random-unitary channels

Let us given a channel $\mathcal{E}$ acting on density matrices ρ defined on the input Hilbert space $\mathscr{H}$. As a consequence of the Stinespring theorem [10], we can write it as follows [11]

$$\mathcal{E}(\rho) = \mathrm{Tr}_a[U(\rho \otimes |0\rangle\langle 0|_a)U^\dagger], \tag{2}$$

namely, as a unitary interaction between the system and an *ancilla* (or *environment*, described by the Hilbert space $\mathscr{H}_a$), followed by a trace over the ancillary degrees of freedom. If the ancilla input state is a pure one—like in Eq. (2)—Gregoratti and Werner [4] proved that, for all possible unitary interactions U in Eq. (2), and for all possible decompositions of the channel $\mathcal{E}$ into pure Kraus representations $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$, there exists a suitable rank-one POVM on the ancilla, let us call it $\{|v_i\rangle\langle v_i|_a\}$, $\sum_i |v_i\rangle\langle v_i|_a = \mathbf{1}_a$, such that

$$E_i \rho E_i^\dagger = \mathrm{Tr}_a[U(\rho \otimes |0\rangle\langle 0|_a)U^\dagger \ (\mathbf{1} \otimes |v_i\rangle\langle v_i|_a)]. \tag{3}$$

As an immediate consequence, if the channel $\mathcal{E}$ admits a random-unitary decomposition as $\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger$, with U_i unitary operators, there exists a rank-one POVM on the ancilla, $\{|\alpha_i\rangle\langle \alpha_i|_a\}$, such that the probability distribution of its outcomes does not depend on the input state ρ, since

$$\mathrm{Tr}[U(\rho \otimes |0\rangle\langle 0|_a)U^\dagger \ (\mathbf{1} \otimes |\alpha_i\rangle\langle \alpha_i|_a)] = \mathrm{Tr}[p_i U_i \rho U_i^\dagger] = p_i, \qquad \forall \rho, \forall i. \tag{4}$$

It is now useful to introduce the channel $\widetilde{\mathcal{E}}$ from density matrices on $\mathscr{H}$ to density matrices on $\mathscr{H}_a$ defined as

$$\widetilde{\mathcal{E}}(\rho) = \mathrm{Tr}_{\mathscr{H}}[U(\rho \otimes |0\rangle\langle 0|_a)U^\dagger]. \tag{5}$$

Since the unitary interaction U is unique up to local isometries on $\mathscr{H}_a$, we can consider such an *ancillary* (or *complementary* [12]) channel as a canonical one. In turn, the channel $\widetilde{\mathcal{E}}$ acting on density matrices, induces a unique *dual* ancillary channel $\widetilde{\mathcal{E}}'$ acting on operators O_a on $\mathscr{H}_a$ as follows

$$\mathrm{Tr}[\widetilde{\mathcal{E}}'(O_a) \ \rho] = \mathrm{Tr}[O_a \ \widetilde{\mathcal{E}}(\rho)]. \tag{6}$$

This is nothing but the Heisenberg picture for the ancillary channel $\widetilde{\mathcal{E}}$. Using this somehow involved notation, we can translate the Gregoratti and Werner theorem stating that a channel $\mathcal{E}$ admits a random-unitary representation (1) *if and only if* there exists a rank-one POVM $\{|\alpha_i\rangle\langle\alpha_i|_a\}$ such that

$$\widetilde{\mathcal{E}}'(|\alpha_i\rangle\langle\alpha_i|_a) = p_i \mathbf{1}_a, \tag{7}$$

for all i, and for some *fixed* probability distribution p_i. In fact

$$p_i = \mathrm{Tr}[\widetilde{\mathcal{E}}(\rho)\ |\alpha_i\rangle\langle\alpha_i|_a] = \mathrm{Tr}[\rho\ \widetilde{\mathcal{E}}'(|\alpha_i\rangle\langle\alpha_i|_a)], \qquad \forall\rho, \forall i. \tag{8}$$

In other words, the POVM $\{|\alpha_i\rangle\langle\alpha_i|_a\}$ is mapped to a classical dice, namely, the POVM $\{p_i\mathbf{1}_a\}$. Notice that the cardinality N of the POVM $\{|\alpha_i\rangle\langle\alpha_i|_a\}$ coincides with the cardinality in the random-unitary decomposition (1).

3 Extremal rank-one POVM's

Let us now suppose that $N > (\dim \mathscr{H}_a)^2$. Then we know that such a POVM is non extremal [13, 14, 15, 16, 17] and it can be convexly decomposed into extremal components

$$|\alpha_i\rangle\langle\alpha_i|_a = \lambda P_i + (1-\lambda)Q_i. \tag{9}$$

(In the above equation we considered a convex combination of just two extremal terms; the general case does not change the conclusions.) Since $\{|\alpha_i\rangle\langle\alpha_i|_a\}$ is rank-one and $0 < \lambda < 1$, the only possibility to satisfy Eq. (9) is that the non-null elements of $\{P_i\}$ and $\{Q_i\}$ are all proportional to the corresponding element of $\{|\alpha_i\rangle\langle\alpha_i|_a\}$. Hence, by linearity, also the non-null elements of $\{P_i\}$ and $\{Q_i\}$ are mapped by $\widetilde{\mathcal{E}}'$ to something proportional to $\mathbf{1}_a$. The normalization is granted by the normalization of the map $\widetilde{\mathcal{E}}'$. This means that at the end we found two other rank-one POVM's, that is $\{P_i\}$ and $\{Q_i\}$, that are both extremal, and hence both with cardinality less or equal to $(\dim \mathscr{H}_a)^2$, achieving two other random-unitary Kraus representations for the channel $\mathcal{E}$.

On the other hand, the normalization condition $\sum_i |\alpha_i\rangle\langle\alpha_i|_a = \mathbf{1}_a$ rules out the possibility that $N < \dim \mathscr{H}_a$. A von Neumann rank-one measurement, with $\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$, achieves the lower bound $N = \dim \mathscr{H}_a$.

4 The result

By now, we showed that a random-unitary channel always admits a random-unitary decomposition (1) involving *at most* $(\dim \mathscr{H}_a)^2$ unitary operators.

We can now tighten this bound by choosing $\dim \mathscr{H}_a$ as small as possible. The smallest[2] achievable $\dim \mathscr{H}_a$ for a given channel $\mathcal{E}$ coincides with the number of Kraus elements in an *orthogonal*—or *canonical*—Kraus representation, that is, $\mathcal{E}(\rho) = \sum_j K_j \rho K_j$ with $\mathrm{Tr}[K_j^\dagger K_l] \propto \delta_{jl}$. Such a number is precisely the rank of the Choi-Jamiołkowski [19, 20] positive operator $R_\mathcal{E}$ in one-to-one correspondence with the channel $\mathcal{E}$ and defined as

$$R_\mathcal{E} = (\mathcal{E} \otimes \mathcal{I})|\Omega\rangle\langle\Omega|, \tag{10}$$

where $\mathcal{I}$ is the identity channel, and $|\Omega\rangle$ is a non normalized ($\|\Omega\|^2 = d$) maximally entangled vector in $\mathscr{H} \otimes \mathscr{H}$. An orthogonal Kraus representation of $\mathcal{E}$ corresponds then to a diagonalization of $R_\mathcal{E}$.

Thus, we have the main result

Theorem *A random unitary channel $\mathcal{E}$ always admits a random-unitary Kraus representation*

$$\mathcal{E}(\rho) = \sum_{i=1}^{K} p_i U_i \rho U_i^\dagger \tag{11}$$

with

$$\mathrm{rank} R_\mathcal{E} \le K \le (\mathrm{rank} R_\mathcal{E})^2. \ \square \tag{12}$$

The bound (12) holds regardless of the dimension d of the input Hilbert space $\mathscr{H}$. It is then reasonable that it fails in accurately describing the peculiar case of qubits ($d = 2$). In fact, it is known that *all* bistochastic qubit channels are actually *Pauli channels* (see, for example, Ref. [9]), that is, they can always be written as (apart from an overall rotation of the whole Bloch sphere)

$$\mathcal{E}(\rho) = \sum_{i=0,x,y,z} p_i \sigma_i \rho \sigma_i, \tag{13}$$

where $\{\sigma_0 \equiv \mathbf{1}, \sigma_x, \sigma_y, \sigma_z\}$ are the usual 2×2 Pauli unitary matrices, and p_i is a probability distribution. Notice that $\mathrm{Tr}[\sigma_i \sigma_j] \propto \delta_{ij}$: This means that the Pauli form of qubit bistochastic channels is a *diagonalization* of the channel itself and the equality

$$K = \mathrm{rank} R_\mathcal{E} \tag{14}$$

[2] See Ref. [18] for a detailed analysis of the ancillary space dimension, that is, the ancillary resources, needed to implement various possible unitary realizations of a given quantum channel.

holds in this case. However, already for $d = 3$ there exist bistochastic channels that *cannot* be diagonalized on unitary operators (for an explicit example, see Ref. [8]). This evidence clearly does not prove our bound to be tight. It nonetheless shows that things, already for $d = 3$, acquire highly non-trivial geometric properties and get much more complicated. In all these cases, the bound given in the Theorem could be tight.

As an immediate consequence of the Theorem, it stems the following

Corollary *The minimum amount of classical information needed to be extracted from the environment in order to perfectly correct a random-unitary channel $\mathcal{E}$ is upper bounded as*

$$H(p_i) \leq 2 \log(\mathrm{rank} R_{\mathcal{E}}).\ \Box \qquad (15)$$

Moreover, since $\mathrm{rank} R_{\mathcal{E}} \leq d^2$, the following quite loose—yet independent of the particular channel—bound holds

$$H(p_i) \leq 4 \log d. \qquad (16)$$

5 Concluding remark

It is noteworthy that we need no more than $(\mathrm{rank} R_{\mathcal{E}})^2$ rank-one POVM elements in order to extract all the "useful" classical information from the ancilla. This is analogous to what happens in the case of optimal *accessible information* extraction: as proved by Davies [21], one never needs more that d^2 rank-one POVM elements in order to extract the maximum achievable accessible information from a d-dimensional system. In the case of accessible information extraction, Davies' bound seems to be tight, in the sense that examples can be constructed in which the maximum information gathering is achieved only by a POVM with maximum number of elements [22]. If the analogy is correct, the bound in the Theorem could be proved to be tight as well, while we expect that the bound given in the Corollary can be refined.

Acknowledgments. This work has been supported by ERATO-SORST Quantum Computation and Information Project. Discussions with P. Perinotti, G. Chiribella, A. Winter, and M. Hayashi are gratefully acknowledged.

References

[1] K. Kraus, *States, Effects, and Operations: Fundamental Notions in Quantum Theory*, Lect. Notes Phys. **190** (Springer-Verlag, Berlin, 1983).

[2] L. J. Landau and R. F. Streater, J. Lin. Alg. Appl. **193**, 107 (1993).

[3] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).

[4] M. Gregoratti and R. F. Werner, J. Mod. Opt. **50**, 915 (2003).

[5] P. Hayden and C. King, Quantum Inf. Comp., **5**, 156 (2005).

[6] J. A. Smolin, F. Verstraete, A. Winter, Phys. Rev. A **72**, 052317 (2005).

[7] A. Winter, preprint `quant-ph/0507045`.

[8] F. Buscemi, G. Chiribella, and G. M. D'Ariano, Phys. Rev. Lett. **95**, 090501 (2005).

[9] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: an Introduction to Quantum Entanglement* (Cambridge University Press, 2006). See, in particular, Chap. 10, Sec. 7.

[10] W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).

[11] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[12] A. S. Holevo, Prob. Th. Appl. **51**, 133 (2006). Available on `quant-ph/0509101`.

[13] A. Fujiwara and H. Nagaoka, IEEE Trans. Inf. Theory **44**, 1071 (1998).

[14] K. R. Parthasarathy, Inf. Dim. Anal. **2**, 557 (1999).

[15] R. Mecozzi, Degree Thesis (Pavia, 2002). In Italian.

[16] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, J. Phys. A: Math. Gen. **38**, 5979 (2005).

[17] M. Hayashi, *Quantum Information: an Introduction* (Springer-Verlag, Berlin, 2006). See Appendix A.4.

[18] F. Buscemi, G. M. D'Ariano, and M. F. Sacchi, Phys. Rev. A **68**, 042113 (2003).

[19] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).

[20] M.-D. Choi, Lin. Alg. Appl. **10**, 285 (1975).

[21] E. B. Davies, IEEE Trans. Inf. Theory **24**, 596 (1978).

[22] P. W. Shor, in *Quantum Communication, Computing, and Measurement 3*, ed. by P. Tombesi and O. Hirota (Kluwer Academic/Plenum Publishers, 2001). Available on `quant-ph/0009077`.